SYSTEM FOR CONTROLLING ACCESS TO RESOURCES IN A STORAGE AREA NETWORK

BACKGROUND OF THE INVENTION

Technical Field

The present invention relates generally to storage area networks, and more particularly, to a method for implementing security management in a

5   storage area network by controlling access to network resources.

Statement of the Problem

A storage area network (SAN) is a dedicated, centrally managed, information infrastructure, which

10   enables interconnection of compute nodes and storage nodes. A storage area network facilitates universal access and sharing of storage resources. SANs are presently being integrated into distributed network environments using Fibre Channel technology (described

15   below). Typically, a SAN utilizes a block-oriented protocol for providing storage to compute nodes, while general purpose networks (GPNs), including local area networks (LANs), wide area networks (WANs) and the Internet, typically implement file-oriented protocols.

20   Storage area networks also differ from general purpose networks in that SANs carry large amounts of data with low latency, and historically have lacked a mechanism for implementing security across the network.

Storage area networks presently typically provide

25   an `everyone (on the network) is trusted' security

1

model because, prior to the availability of Fibre Channel, SANs had a distance limitation on the order of tens of meters. Therefore, compute node operating system (O/S) behavior in existing storage area

5 networks has been in accordance with the distance constraint, i.e., there has been little relatively storage resource sharing among different compute node and each compute node often has dedicated data storage.

10 Compute nodes on SANs are often also server nodes of a GPN. In these networked systems, the SAN is often implemented with separate, high-speed, network hardware from that of the GPN so as to offload the data from the GPN, thereby increasing GPN and

15 effective CPU performance. Such separation is often desirable because effective compute node CPU performance is often limited by the available bandwidth between compute and storage nodes, and because the bandwidth required between compute and

20 storage nodes often far exceeds all other network traffic affecting the same compute nodes.

Development of storage area networks has been motivated by the need to manage and share the dramatically increasing volume of business data, and

25 to mitigate its effect on GPN performance. Using Fibre Channel connections, SANs can provide high-speed compute node to/from storage node, and storage node to storage node, communications at distances that allow remote workstation and server compute nodes to easily

30 access large shared data storage pools.

Using SAN technology, management of storage systems can be more easily centralized than with alternative technologies, and data backup is facilitated. Both factors act to increase overall

2

system efficiency. The large distances allowed by Fibre Channel SAN technology make it easier to deploy remote disaster recovery sites than with prior technology.

5   A Fibre Channel SAN can be local, or can now be extended over large geographic distances. The SAN can be viewed as an extension to the storage bus concept that enables storage devices and servers to be interconnected using similar elements as in local area
10   networks (LANs) and wide area networks (WANs): routers, hubs, switches and gateways.

       Fibre Channel is presently considered to be the architecture on which most future SAN implementations will be built. Fibre Channel is a technology standard
15   that allows data to be transferred from one network node to another at very high speeds. This standard is backed by a consortium of industry vendors and has been accredited by the American National Standards Institute (ANSI). The word *Fibre* in Fibre Channel is
20   spelled differently than *"fiber"* to indicate that the interconnections between nodes are not necessarily based on fiber optics, but can also use copper cables. Fibre Channel is, in essence, a high performance serial link supporting its own, as well as higher
25   level protocols such as the FDDI, SCSI, HIPPI and IPI. SAN configurations may incorporate the FIGURE 5 protocol encapsualted within Fibre Channel frames.

       Data integrity is an important issue in storage area network technology, since multiple compute nodes
30   employing diverse types of operating systems could coexist within the SAN, and some operating systems do not gracefully share access to the same storage devices with other operating systems. Some operating systems do not even gracefully share access to storage

devices among multiple compute nodes even if each node runs the same O/S. Because of this, conflicts can occur that can have damaging results. These conflicts may include file and record lock conflicts, overwrites of home blocks on previously initialized disks, reservations taken out on disks which a compute node should not have access to, improper reformatting, overwriting of files, or other maloperation.

Presently, many current SAN implementations rely on limits on access to the physical wiring for security purposes. As SANs become larger and more geographically dispersed, a security scheme is required which will provide SAN-wide security in order to prevent conflicts over the entire network.

One security mechanism presently being implemented is a partitioning approach called 'zoning', or effectively partitioning at the 'wire' level of the SAN. Various levels of 'zoning' may be used to restrict the any-to-any access by limiting compute node attachment to specific storage nodes. Zoning is often implemented in Fibre Channel switches, such as those available from Brocade Communications Systems, Inc. These switches can be programmed to filter Fibre Channel frames according to their source and destination identifiers, thereby restricting SAN communications to those among authorized nodes and node pairs. Multiple such switches may be incorporated into a switched fabric that appears to each node as a larger, potentially geographically dispersed, switch.

It is known that storage nodes of a SAN could be RAID (Redundant Array of Independent Disks) controllers. RAID controllers are also known as array controllers since they are typically operable to

4

present storage to a SAN with or without operating in
redundancy modes. The RAID controllers could be
configured to serve multiple logical units of storage.
Each logical unit may represent a physical disk or
5   tape drive, or be formed from part of one, all of one,
or a combination of several, disk drives with or
without redundancy. Each logical unit is a storage
resource intended for use by a set of one or more
compute nodes, where the sets of intended compute
10   nodes for each logical unit may differ. Zoning as
currently implemented typically restricts
communications on a node basis, not a logical unit
basis.

Resource providers of a SAN include storage nodes
15   as well as any other node configured to provide
resources to the SAN. Similarly, resource users of a
SAN include compute nodes as well as any other node
configured to use resources available on the SAN. For
example, but not by way of limitation, a storage node
20   having a data backup device and a disk device can be
simultaneously both a resource provider - providing
disk LUNs to the SAN - and a resource user - accessing
disk resources of other storage nodes to backup data.
Similarly, a storage node having disk devices could be
25   a resource provider - providing disk LUNs to the SAN -
and a resource user - transmitting data changes to a
second resource provider to maintain a mirrored
dataset.

Therefore, security and access control needs to
30   be improved to guarantee data integrity by preventing
conflicts. It is also desirable that the security and
access controls be capable of management from a single
network management point.

The Fibre Channel specification does not include a specific mechanism for managing security-related issues, and there is presently no commonly available solution to the above-described problems of providing secure

5   access to shared SAN resources.

Solution to the Problem

The present invention overcomes the aforementioned problems of the prior art and achieves an advance in the field by providing a method for
10   controlling and managing resources on storage area networks (SANs). The method is applicable to a wide range of storage area networks, including large scale storage area networks.

The present system provides a table-driven
15   mechanism whereby resources at the LUN level on the SAN can be allocated to specific resource users. Unless authorized to do so, a compute node or other resource user on the SAN is not allowed to access a particular resource.

20   Two tables are maintained at each resource provider to control storage resources. These include a table for 'approved' storage users (identified by port WWN and node WWN) and the approved resources (typically logical units, or LUNs) to which they have
25   been granted access, and one for 'not-yet-approved' storage users (identified by port WWN and node WWN) that can see, but have not been granted resource access approval by, the resource provider. The table of approved storage resource users may be stored in
30   non-volatile memory of each storage RAID controller so that only initial setup of the table is required. If the table of approved hosts/resources is stored in non-volatile memory, the information therein is

6

available after any event that would require a
resource user to re-poll for resources, such as a re-
boot operation. The table of 'not-yet-approved'
resource users is provided to allow a system
5   administrator to copy information therefrom to the
'approved' table rather than having to enter the
information by hand. This 'not-yet-approved' table
may, but need not, be stored in volatile memory, which
is often less expensive than non-volatile memory.

10   To facilitate operation of the present system, a
SAN management interface is made available to a system
manager. The use of the SAN management interface is
password protected by the use of known network
protocols and standards. At this management
15   interface, the system manager creates or adds to the
'approved' table of each resource provider, matching
resource users with access to specific LUNs. Once
permission has been granted, information maintained in
the 'approved' table facilitates subsequent access by
20   the specified resource user or users.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a diagram showing the present system
in the context of a storage area network;

Figure 2 is a flowchart illustrating, at a high
25   level, steps which may be performed when a host
attempts to access a resource provider;

Figure 2A is a flowchart illustrating, at a high
level, steps which may be performed when an
administrator reconfigures the system to change an
30   'approved entity' table;

Figure 2B is a flowchart illustrating, at a high
level, steps which may be performed when an

7

administrator reconfigures the system to change an 'approved entity' table;

Figure 3 is a diagram showing exemplary contents of the 'approved' table; and

Figure 4 is a diagram showing exemplary contents of the 'not-yet-approved' table.

Figure 5 is a diagram showing exemplary contents of the Report LUNs command when a RAID controller or other resource provider informs a host 103/105 of the resources that it may access. This is a more detailed description of the access map of figure 3 item 309.

Figure 6 is a diagram showing exemplary contents of the table returned by the controller to the management interface describing the available storage.

DETAILED DESCRIPTION OF THE EMBODIMENTS

An embodiment of the present system 100 (figure 1) is illustrated in the context of an exemplary storage area network (SAN). The storage area network includes a switch 104 which interconnects a plurality of resource providers, here RAID controllers 101 (shown as controllers 101(1)—101(n) in Figure 1) with a plurality of resource users, here compute nodes including servers 103 and workstations 105. A typical storage area network uses Fibre Channel technology for intercommunication between the entities connected to the network.

In the exemplary embodiment of Figure 1, each compute node, such as host processor (or server) 103, workstation 105, or other resource user (hereinafter collectively referred to as "host 103/105") is coupled to Fibre Channel switch 104 via an adapter 17,

8

appropriate cabling, and other network components.
Switch 104 may also be a switched fabric, comprising
multiple switches in communication with each other.
Also connected into the SAN is management station 106,
5   which comprises a processor, associated memory, an
input device, and a display device. Management
station 106 (described in detail below) provides the
capability to manually enter configuration data into a
given RAID controller 101(*) in a secure manner [the
10   asterisk following a reference number is a 'wild card'
character indicative of any one of the entities having
the associated reference number].

     Each RAID controller 101 is connected to a disk
array 102 containing a plurality of disk storage
15   devices. Each RAID controller 101 contains volatile
memory (e.g., RAM) 13 and non-volatile memory 14.
Volatile memory 13 in each controller 101 contains a
table 15, the 'not-yet-approved entity table') in
which is stored SAN hosts 103/105 that presently do
20   not have any 'approved' resources. Non-volatile
memory 14 in each controller 101 contains an 'approved
entity' table 16 in which is stored:

          (1) SAN entities with 'approved' resources,

          (2) a list of those resources, and

25   (3) other host 103/105 configuration information.

The function of each of these tables 15/16 is
described in detail below. The term "resources", as
used herein, refers to data storage devices such as
magnetic or optical disk and tape drives, and may for
30   some devices refer more specifically to data storage
areas, such as logical units (LUNs) such as may exist
on disk storage devices in a disk array 102.

9

Figure 2 is a flowchart illustrating, at a high level, steps which may be performed in practicing one embodiment of the method of the present invention. As shown in Figure 2, at step 205, a resource user, such as a host computer or server 103, or a workstation 105, 'discovers' a potential resource provider 101 on the storage area network, in response to notification that the particular resource is available on the SAN. There are a number of possible ways in which resources can be discovered; for example, a host adapter 17 may poll switch 104, or the host 103/105 may subscribe to a service that provides notification through switch 104 of new resources (i.e., disk storage arrays) 101 that become available on the SAN. In the presently-described embodiment, potential resource users (hosts) 103/105(*) initiate communication to a resource provider (an RAID controller) 101 to discover what resources are available.

At step 210, a host 103/105(*) logs-in to the controller 101 and presents its node WWN (World-Wide Name, a unique identifier) and port WWN to the discovered RAID controller 101. In doing so, the host 103/105(*) provides the source identifier S_ID that will identify communications from that host. RAID controller 101 or other resource provider examines the 'approved entity' table 16 in the controller's non-volatile memory 14, at step 215, to determine whether resources 102(*) are available to the requesting host 103/105(*) . The verification made at step 215 includes verifying that the node WWN and port WWN supplied by host 103/105 match corresponding entries in the 'approved entity' table 16, described in detail below. Alternatively, this verification may include only checking either the node or port WWN, or

10

verifying other host identification information previously established and stored in table 16.

If authorized, the host may then send commands to the resource. For every command sent, the host provides a source identification (S_ID) that identifies frames associated with that host.

Exemplary contents of the 'approved entity' table 16 are illustrated in Figure 3. As shown in Figure 3, the 'approved entity' table 16 contains information indicative of the specific resources 102 allocated to each resource user (host) 103/105. It is advantageous to store 'approved entity' table 16 in non-volatile memory so that it need not be set up repeatedly and remains intact after events, such as power failure, RAID controller re-boot, etc., that would cause resource users 103/105 to re-poll for resources. Approved entity table 16 makes RAID controller resources 102 accessible to resources consumers 103/105 by providing information such as that included in the following fields indicated by the reference number appearing in Figure 3.:

301 Host port WWN;

302 Host Node WWN;

303 Host symbolic name;

304 Unit offset _host _relative address of presented LUN resource;

305 Pending unit attentions _information about the state of the storage that is held until the host issues a command to a given device;

307 S_ID an identifier on commands indicating the source host application. While not necessarily

11

unique throughout the world to the host like the WWN,
S_ID is typically unique to the host within the SAN;

308   Persistent Reserve Information _ a SCSI
standard host support mechanism that allows hosts to
manage locking and resource allocation dynamically;

309   LUN access map _indicates which LUNs this
host (port WWN and node WWN) has access to; and

310   Host Mode _controller tailors specific
behaviors to what the host is looking for; for
example, this field might indicate that this is an NT
host, or an IBM host.

Fields 301—310 (listed above) are entered into
'approved entity' table 16 in step 245, described
below. In an exemplary embodiment of the present
system, 'approved entity' table 16 may contain up to
256 entries, each of which is slightly larger than 4K
bytes, primarily due to the size of Persistent Reserve
Information field 308, which is itself a 4K byte
length field. In this embodiment, approximately 1MB
of non-volatile memory is required to contain 256
entries. It is anticipated that other embodiments may
permit the table to contain other maximum numbers of
entries, including numbers of entries greater than
256.

When a host 103/105 presents its node WWN and
port WWN to the resource provider or RAID controller
101, the resource provider checks its 'approved
entity' table 16 in the controller's non-volatile
memory. If, at step 215, the host port WWN and node
WWN are listed in table 16, then at step 220, resource
provider 101 returns the LUN access map 309 for the
entry to the host. LUN access map 309 indicates which

12

resources 102 managed by resource provider (RAID
controller) 101 are to be made accessible to a given
host 103/105, and therefore, only a single entry in
table 16 is required for access to multiple resources
5   by a single host.

The LUN access map 309 is provided to the host
through a SCSI "Report LUNs" command embedded in fibre
channel frames. The response to this command is
formatted according to Figure 5, and reports those
10 LUNs allocated to that host. LUNs not yet allocated
to that host are not presented to it.

If, at step 215, it is determined that the
requested connection is not stored in the 'approved
entity' table 16, then at step 225, the Host port WWN
15 and Host Node WWN from the log-in request is placed in
the 'not-yet-approved entity' table 15 and no
resources are provided.

It is anticipated that the presently described
embodiment may be operable with or without zoning
20 implemented in the switch or switched fabric 104. In
the event that zoning is not implemented, the 'not-
yet-approved' table of each resource provider will
tend to acquire entries corresponding to all resource
users, or hosts, in the SAN. In the event that zoning
25 is implemented, the 'not-yet-approved' table of each
resource provider will tend to acquire entries
corresponding to all resource users, or hosts, in the
SAN that have been zoned such that they can
communicate with that resource provider.

30   Figure 4 is a diagram showing exemplary contents
of the 'not-yet-approved entity' table 15. As shown
in Figure 4, each entry in table 15 contains two or
more fields, Host port WWN 401 and Host Node WWN 402.

13

In an exemplary embodiment of the present system, 'not-yet-approved entity' table 15 may contain up to 1024 entries, or large enough to store the WWN information for every resource in the SAN. In this

5   embodiment, each of the entries 401/402 in table 15 is 8 bytes in length, yielding a total size of approximately 16K bytes. It is anticipated that other embodiments may permit larger numbers of entries, and that additional information, such as a symbolic host

10   name 404, may optionally be stored in this table. The entries in 'not-yet-approved entity' table 15 make it easier for management station 106 user to create the 'approved entity' table 16 (described in step 245, below). Table 15 is stored in volatile memory, since

15   it does not need to be preserved across power failures or re-boots by the associated RAID controller 101.

In order to allocate resources to a host 103/105 (as identified by port WWN and node WWN), a system manager or SAN administrator logs in 240 (Figure 2A)

20   to management station 106 and selects a resource provider, here RAID controller 101(1) to be managed. Login to the management server 106, and from the management server to the selected RAID controller 101(1), are password protected to prevent unauthorized

25   or accidental changes to resource allocations.

A list (Figure 6) of available resources 102, the 'approved entity' table 16, and the 'not-yet-approved' entity table 15 are transmitted from RAID controller 101(1) to the management station 106. A management

30   client of management station 106 formats and displays information from the tables and the list of available resources. After viewing the available resources, the administrator selects a host by selecting a port WWN, a node WWN, or a symbolic hostname from either the

14

table 'not-yet-approved entity' 15 or the 'approved entity' table 16. Selection of a host from either table may be accomplished through a mouse-click or other selection method.

5   Once the administrator has selected a host, the administrator then selects, from the list of available resources, the available resource or resources to be made available to, or remove from availability to, the selected host. As the administrator makes selections

10   indicating desired changes, a record of the desired changes is made in memory of the management station 106.

       Management station 106 then conveys the record of the desired changes to RAID controller 101(1) in a

15   secure manner. RAID controller 101(1) in turn, then performs the following actions to update the 'approved entity' table as per the desired updates. For example, updating the 'approved entity' table 16 to allow access to an additional resource may involve the

20   steps

       (1) if the designated port WWN and node WWN combination appears in 'not-yet-approved entity' table 15 then

       (a) remove it and create a new entry in

25   table 16;

       (b) identify the new entry with port WWN and node WWN; and

       (c) create a LUN access map 309 and initialize remaining fields of the table;

15

(2) else, if port WWN and node WWN already appears in 'approved entity' table 16, then add the designated LUN to the LUN access map 309.

At step 240 (Figure 2A), the SAN administrator logs into the management station 106 using a standard LAN/WAN password. This may be performed directly at the management station or over a separate TCP/IP or other GPN network interface that has proven security features. The purpose, in this case is to present new storage resources to a new host or allocate more resources to an established (in table 16) host. The administrator then designates a particular RAID controller 101(1) to undergo allocation changes.

At step 245, the management station 106 requests the current 'approved entity' table 16, as well as the 'not yet approved entity' table 15 and a resource list, which are returned by the selected RAID controller 101(1). At the management station 106 the SAN administrator is able to view the list of current resources Figure 6 and then associates an available resource with a host at step 250. The management station 106 then transmits 255 the association to the resource provider. The management station 106 may be used additionally to configure host preferences 310, R/W access 306, and unit offset information 304, and, optionally, host symbolic name 303.

Upon receiving the association, the RAID controller 101 checks 260 the host entry in its 'not-yet-approved entity' table 15. If it is not found in this table, a check is performed to see if the designated host is already present in the 'approved entity' table 16; if it is not present it is added by moving 285 the associated entry from the 'not yet approved' table 15 to the 'approved entity' table 16.

16

If the designated host was already present in the
'approved entity' table 16, a check is made 265 to
determine if the designated resource is already
associated with the host in the 'approved entity'
5   table.   If so, a further check 267 is made to verify
that the administrator is de-associating the resource,
and an warning is issued if not.   De-association, or
release of resources, is performed by deleting 270 the
available resource entry from the LUN map of the host
10   entry.   If 275 there are no more active LUNs in the
LUN map for that host entry the host entry is moved
280 from the 'approved entity' table to the 'not-yet-
approved entity' table.

When adding an association between a host and a
15   resource, the resource identification is added 290 to
the LUN map of the host entry in the 'approved entity'
table 16.

The selected resource provider 101 saves the
altered 'approved entity' table in NV memory 14 in the
20   selected RAID controller 101(1); and the altered 'not-
yet-approved entity' table in volatile memory.
Finally, the RAID controller 101(1) notifies 295 the
indicated host that the RAID controller may have
changed its resources.

25   Finally, at step 295, when a change in resources
is presented to a given host 103/105(*) by RAID
controller 101(*), host discovery of SAN resources is
initiated, so that hosts 103/105 may be made aware of
the available resources 102 on the storage area
30   network.

Once one or more hosts has discovered and has
been granted access to resources, following frames
attempting to read or write resources are filtered to

17

eliminate those that come from hosts that have not logged-in to the resources. For example, any frame attempting to read or write a resource that has a source identification (S_ID) not associated with a host in the 'approved entity' table will be rejected.

A system for selectively presenting logical units (LUNs) to host computing systems, disclosed in co-pending U.S. Patent Application entitled "System and Method for Selectively Presenting Logical Storage Units to Multiple Host Operating Systems in a Networked Computing System", Serial No. 09/312,944, filed May 17, 1999 (the '944 patent application), is incorporated herein by reference. The system disclosed therein includes an RAID controller or other resource provider 101 for controlling and coordinating the operations of persistent storage devices 102; a memory 14 accessible by the RAID controller; and a configuration table (not shown in Figure 1), stored in the memory 14, containing entries governing the interactions between the storage devices 102 and the one or more host computing systems 103/105. The configuration table stores information relating to the host connection through which a particular logical unit on a storage device 102 should communicate; any LUN number offsets employed by the logical unit to communicate with a particular host 105/105; and any special "host modes" (see table 16, field 310) employed by an RAID controller 101 in communicating with a particular host. It is to be noted that the '944 patent application discloses neither a table, nor management station 106 with password protection. Furthermore, by default, in the system disclosed in the '944 patent application, all resources are shown to any host that logs in; thus there is no mechanism

18

disclosed for preventing any host from logging in to any resource on the network.

While preferred embodiments of the present invention have been shown in the drawings and described above, it will be apparent to one skilled in the art that various embodiments of the present invention are possible. For example, the specific configuration of tables 15 and 16 as well as the particular entities in the storage area network described above should not be construed as limited to the specific embodiments described herein. Modification may be made to these and other specific elements of the invention without departing from its spirit and scope as expressed in the following claims.